

SPHERICAL AVERAGES OF SIEGEL TRANSFORMS FOR HIGHER RANK DIAGONAL ACTIONS AND APPLICATIONS

JAYADEV S. ATHREYA, ANISH GHOSH, AND JIMMY TSENG

ABSTRACT. We investigate the geometry of approximates in multiplicative Diophantine approximation. Our main tool is a new averaging result for Siegel transforms on the space of unimodular lattices in \mathbb{R}^n which is of independent interest.

CONTENTS

1. Introduction	1
2. Equidistribution on the space of lattices	6
3. Proof of Theorem 2.3	8
4. Appendix	13
5. Concluding remarks and generalisations	14
References	14

1. INTRODUCTION

The main result in the present paper is a new averaging theorem for Siegel transforms on the homogeneous space $\mathrm{SL}_n(\mathbb{R})/\mathrm{SL}_n(\mathbb{Z})$. Such results have found several applications in number theory and indeed our motivation is to investigate the distribution of approximates in certain foundational results in Diophantine approximation. In [1], we studied the phenomenon of *spiraling* of approximates in Dirichlet's theorem and obtained a number of distribution results for approximates. In the present paper, we continue our investigations in this subject and present a new multi parameter averaging result for Siegel transforms and as a consequence, obtain new results on the geometry of approximates in *multiplicative* and *weighted* Diophantine approximation. We briefly recall the setup in [1] and then state our main results. The bulk of the paper is concerned with the proof of Theorem 2.3, our result on averages of Siegel transforms. The general principle that equidistribution of spherical averages implies distribution results for approximates applies in a wide variety of situations. In the final section, we briefly survey some such situations.

1.1. Dirichlet's theorem and spiraling. Let $\alpha_{ij}, 1 \leq i \leq m, 1 \leq j \leq n$ be real numbers and $Q > 1$. Then Dirichlet's theorem in Diophantine approximation states that there exist integers $q_1, \dots, q_m, p_1, \dots, p_n$ such that

$$(1.1) \quad 1 \leq \max\{|q_1|, \dots, |q_m|\} \leq Q$$

2000 *Mathematics Subject Classification.* 37A17, 11K60, 11J70.

Key words and phrases. Diophantine approximation, equidistribution, Siegel transforms.

J.S.A. partially supported by NSF grant DMS 1069153, and NSF grants DMS 1107452, 1107263, 1107367 "RNMS: GEometric structures And Representation varieties" (the GEAR Network).

A.G. partially supported by an ISF-UGC grant.

J.T. acknowledges the research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n. 291147.

and

$$(1.2) \quad \max_{1 \leq i \leq n} |\alpha_{i1}q_1 + \cdots + \alpha_{im}q_m - p_i| \leq Q^{-m/n}.$$

In our earlier work [1], we studied the problem of *spiraling* of the approximates appearing in Dirichlet's theorem and showed as a consequence, that on average, the directions of approximates spiral in a uniformly distributed fashion on the unit sphere of one lower dimension. In fact, the problem can be recast as a special case of a more general equidistribution result in the space of lattices. As far as we are aware, this is the only work addressing the natural question of how the approximates appearing in Dirichlet's theorem are distributed. In [1], we considered vectors rather than linear forms although the proof goes through for linear forms with very minor modifications.

Given $\mathbf{x} \in \mathbb{R}^d$, we form the associated unimodular lattice in \mathbb{R}^{d+1}

$$\Lambda_{\mathbf{x}} := \begin{pmatrix} \text{Id}_d & \mathbf{x} \\ 0 & 1 \end{pmatrix} \mathbb{Z}^{d+1} = \left\{ \begin{pmatrix} q\mathbf{x} - \mathbf{p} \\ q \end{pmatrix} : \mathbf{p} \in \mathbb{Z}^d, q \in \mathbb{Z} \right\}.$$

Then we can view the approximates (\mathbf{p}, q) of \mathbf{x} appearing in Dirichlet's Theorem as points of the lattice $\Lambda_{\mathbf{x}}$ in the region

$$(1.3) \quad R := \left\{ \mathbf{v} = \begin{pmatrix} \mathbf{v}_1 \\ v_2 \end{pmatrix} \in \mathbb{R}^d \times \mathbb{R} : \|\mathbf{v}_1\| |v_2|^{1/d} \leq 1 \right\}.$$

The set R is a thinning region around the v_2 -axis, and the following sets are used to study the distribution of lattice approximates in R . Let

$$(1.4) \quad R_{\epsilon, T} := \{ \mathbf{v} \in R : \epsilon T \leq v_2 \leq T \}$$

and, for a subset A of \mathbb{S}^{d-1} with zero measure boundary,

$$(1.5) \quad R_{A, \epsilon, T} := \left\{ \mathbf{v} \in R_{\epsilon, T} : \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} \in A \right\}.$$

For a unimodular lattice Λ , define

$$N(\Lambda, \epsilon, T) = \# \{ \Lambda \cap R_{\epsilon, T} \}$$

and

$$N(\Lambda, A, \epsilon, T) = \# \{ \Lambda \cap R_{A, \epsilon, T} \}.$$

Let dk denote Haar measure on $K := K_{d+1} := \text{SO}_{d+1}(\mathbb{R})$, and let $X_{d+1} := \text{SL}_{d+1}(\mathbb{R}) / \text{SL}_{d+1}(\mathbb{Z})$. In [1], we proved

Theorem 1.1. *For every $\Lambda \in X_{d+1}$, $A \subset \mathbb{S}^{d-1}$ as above, and for every $\epsilon > 0$,*

$$(1.6) \quad \lim_{T \rightarrow \infty} \frac{\int_K N(k^{-1}\Lambda, A, \epsilon, T) dk}{\int_K N(k^{-1}\Lambda, \epsilon, T) dk} = \text{vol}(A).$$

The main tool in proving Theorem 1.1 is an equidistribution result for spherical averages. Given a lattice Λ in \mathbb{R}^{d+1} and a bounded Riemann-integrable function f with compact support on \mathbb{R}^{d+1} , denote by \hat{f} its *Siegel transform*:

$$\hat{f}(\Lambda) := \sum_{\mathbf{v} \in \Lambda \setminus \{0\}} f(\mathbf{v}).$$

Then

Theorem 1.2. *Let f be a bounded Riemann-integrable function of compact support on \mathbb{R}^{d+1} . Then for any $\Lambda \in X_{d+1}$,*

$$\lim_{t \rightarrow \infty} \int_{K_{d+1}} \hat{f}(g_t k \Lambda) dk = \int_{X_{d+1}} \hat{f} d\mu.$$

1.2. Multiplicative and weighted variants. Dirichlet's theorem lends itself to several interesting generalisations. Here is a *multiplicative* analogue which can be proved using either Dirichlet's original approach or Minkowski's geometry of numbers. With notation as above, there exist integers $q_1, \dots, q_m, p_1, \dots, p_n$ such that

$$(1.7) \quad \left(\prod_{1 \leq j \leq m} \max\{1, |q_j|\} \right)^{1/m} \leq Q$$

and

$$(1.8) \quad \left(\prod_{1 \leq i \leq n} |\alpha_{i1}q_1 + \dots + \alpha_{im}q_m - p_i| \right)^{1/n} \leq Q^{-m/n}.$$

As a corollary, it follows that there are infinitely many q_1, \dots, q_m such that

$$(1.9) \quad \left(\prod_{1 \leq i \leq n} |\alpha_{i1}q_1 + \dots + \alpha_{im}q_m - p_i| \right) \leq \left(\prod_{1 \leq j \leq m} \max\{1, |q_j|\} \right)^{-1}$$

for some p_1, \dots, p_n .

The study of Diophantine inequalities using the multiplicative “norm” as above instead of the supremum norm is referred to as *multiplicative Diophantine approximation*. This subject is considered more difficult and is much less understood in comparison to its standard counterpart. For instance, arguably the most emblematic open problem in metric Diophantine approximation namely the Littlewood conjecture, is a problem in this genre. We refer the reader to the nice survey [3] by Bugeaud for an overview of the theory. There have been several important advances recently, several arising from applications of homogeneous dynamics to number theory. We mention the work of Kleinbock and Margulis [11] settling the Baker-Sprindzhuk conjecture as well as the work of Einsiedler-Katok-Lindenstrauss making dramatic progress towards Littlewood's conjecture.

Another variation of Diophantine approximation is developed as follows. Let $\alpha_{ij}, 1 \leq i \leq m, 1 \leq j \leq n$ be real numbers and let $\mathbf{r} = (r_1, \dots, r_n) \in \mathbb{R}^n$ and $\mathbf{s} = (s_1, \dots, s_m) \in \mathbb{R}^m$ be probability vectors. Recall that a *probability vector* has nonnegative real components, the sum of which is equal to 1. Then a weighted version of Dirichlet's theorem states that there exist infinitely many integers q_1, \dots, q_m such that

$$(1.10) \quad \max_{1 \leq i \leq n} |\alpha_{i1}q_1 + \dots + \alpha_{im}q_m - p_i|^{1/r_i} \leq \max_{1 \leq j \leq m} |q_j|^{1/s_j}$$

for some p_1, \dots, p_n . The subject of *weighted* Diophantine approximation has also witnessed significant progress of late. We refer the reader to the works of Kleinbock and Weiss, [13, 14] as well as the resolution of Schmidt's conjecture on weighted badly approximable vectors due to Badziahin-Pollington-Velani [2].

1.3. Spiraling. In this paper, we study the distribution of approximates in the multiplicative setting as well as the setting of Diophantine approximation with weights. Again, as far as we are aware, these are the first results of their kind. While our strategy remains the same as in [1], our main tool, an equidistribution theorem for Siegel transforms on homogeneous spaces (Theorem 2.2) is new and new inputs are required for the proof. Equidistribution results of this kind have found many applications (cf. [11], [8], [16, 15]) in number theory. We hope our result will be of interest to both dynamicists as well as number theorists.

1.4. **The setup.** Let $\ell \geq 1$ be an integer. Define functions $\mathbb{R}^\ell \rightarrow \mathbb{R}_{\geq 0}$ as follows:

$$\|\mathbf{v}\|_{\mathbf{p}} := \max_{i=1, \dots, \ell} |v_i|^{1/p_i} \quad \text{and} \quad \|\mathbf{v}\|_{\text{pr}} := \prod_{i=1}^{\ell} |v_i|$$

where $\mathbf{p} \in \mathbb{R}^\ell$ is a probability vector. Let $m, n \geq 1$ be integers and $d := m + n$. Let e_1, \dots, e_m be the standard basis for \mathbb{R}^m and e_1, \dots, e_d be the standard basis for $\mathbb{R}^m \times \mathbb{R}^n = \mathbb{R}^d$. Fix probability vectors $\mathbf{r} \in \mathbb{R}^m$ and $\mathbf{s} \in \mathbb{R}^n$, these vectors are also referred to as *weights* in the literature. Let

$$g_t^{(\mathbf{r})} := \text{diag}(e^{r_1 t}, \dots, e^{r_m t}) \in \text{GL}_m(\mathbb{R}),$$

and let \mathbb{S}^{m-1} denote the $m-1$ dimensional unit sphere centered at the origin. For a subset \tilde{A} of \mathbb{S}^{m-1} , the union of all rays in \mathbb{R}^m through each point of \tilde{A} is called the *cone in \mathbb{R}^m through \tilde{A}* and denoted by $\mathcal{C}\tilde{A}$. The region of interest for Diophantine approximation with weights is

$$R := R^{(\mathbf{r}, \mathbf{s})} := \left\{ \mathbf{v} = \begin{pmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{pmatrix} \in \mathbb{R}^m \times \mathbb{R}^n : 0 < \|\mathbf{v}_1\|_{\mathbf{r}} \|\mathbf{v}_2\|_{\mathbf{s}} \leq 1 \right\}.$$

Fix an $0 < \epsilon < 1$, $T > 0$, and a subset A of \mathbb{S}^{m-1} with zero measure boundary. The subsets that concern us, in particular, are

$$R_{\epsilon, T} := \{ \mathbf{v} \in R : \epsilon T \leq \|\mathbf{v}_2\|_{\mathbf{s}} \leq T \} \quad \text{and} \quad R_{A, \epsilon, T} := \left\{ \mathbf{v} \in R_{\epsilon, T} : \mathbf{v}_1 \in g_{-\log(T)}^{(\mathbf{r})}(\mathcal{C}A) \right\}.$$

The subset $R_{\epsilon, T}$ is analogous to the subset above which played a role in [1]. Indeed if we consider the special case of \mathbf{r} equal to $(1/m, \dots, 1/m)$, then the set $R_{A, \epsilon, T}$ is equal to $\left\{ \mathbf{v} \in R_{\epsilon, T} : \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|_2} \in A \right\}$, which was considered in [1]. The reason that our formulation in terms of cones is the appropriate generalization is as follows. Let us again consider an arbitrary \mathbf{r} . Consider the slices of R given by the equations

$$\|\mathbf{v}_1\|_{\mathbf{r}} = 1/p$$

for a real number $p > 1$. To map the slice given by p to the one given by $p' \geq p$, apply the contracting (and, in general, nonuniformly contracting) automorphism $g_{\log(p) - \log(p')}^{(\mathbf{r})}$ to the slice. Now $g_{-t}^{(\mathbf{r})}$ takes \mathbb{S}^{m-1} into ellipsoids, whose eccentricities are increasing as t increases. It is reasonable that the distribution of directions respects the action of $g_{-t}^{(\mathbf{r})}$ —that this holds is the content of our result, Theorem 1.3.

The regions of interest for multiplicative Diophantine approximation are

$$P := \left\{ \mathbf{v} = \begin{pmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{pmatrix} \in \mathbb{R}^m \times \mathbb{R}^n : 0 < \|\mathbf{v}_1\|_{\text{pr}} \|\mathbf{v}_2\|_{\text{pr}} \leq 1 \right\},$$

$$P_{\epsilon, T} := \{ \mathbf{v} \in P : \epsilon T \leq \|\mathbf{v}_2\|_{\text{pr}} \leq T \} \quad \text{and} \quad P_{A, \epsilon, T} := \left\{ \mathbf{v} \in P_{\epsilon, T} : \mathbf{v}_1 \in g_{-\log(T)}^{(\mathbf{r})}(\mathcal{C}A) \right\}.$$

The region P is sometimes referred to as a *star body*. For the special case of \mathbf{r} equal to $(1/m, \dots, 1/m)$, the set $P_{A, \epsilon, T}$ is equal to $\left\{ \mathbf{v} \in P_{\epsilon, T} : \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|_2} \in A \right\}$. Now, unlike for Diophantine approximation with weights, the m -volume of $P_{1,1}$ is infinite. Let \mathbb{P}_i denote the coordinate codimension-one hyperplane in \mathbb{R}^m normal to e_i . Then

$$\mathbb{P}_i \cap \mathbb{S}^{m-1} =: \mathbb{S}_i$$

are *great spheres* of \mathbb{S}^{m-1} ; namely, $\mathbb{S}_i = k_i \mathbb{S}^{m-2}$ for some $k_i \in \text{SO}_m(\mathbb{R})$. For any $\delta > 0$, let

$$\mathbb{S}_i^{(\delta)} := \mathbb{P}_i \times [-\delta, \delta] \cap \mathbb{S}^{m-1}$$

denote the δ -thickening of \mathbb{S}_i on \mathbb{S}^{m-1} . By elementary calculus, it is easy to see that the \mathbb{P}_i point in the directions in which $P_{1,1}$ has regions with infinite volume (see also the Appendix). Radially projecting $P_{1,1}$ onto

$$\mathcal{S} := \mathcal{S}(\delta) := \mathbb{S}^{m-1} \setminus \bigcup_{i=1}^m \mathbb{S}_i^{(\delta)}$$

it is easy to see that $\mathcal{CS} \cap P_{1,1}$ has finite m -volume for every $\delta > 0$. We also note that the $g_{-t}^{(\mathbf{r})}$ -action contracts slices of P in the same way as it does R and that it preserves each of the coordinate planes: $g_{-t}^{(\mathbf{r})}(\mathbb{P}_i) = \mathbb{P}_i$, and consequently, the action of $g_{-t}^{(\mathbf{r})}$ on $\mathcal{CS} \cap P_{1,1}$ keeps the m -volume finite. By continuity in t , the m -volume of slices between $\epsilon T \leq t \leq T$ has a maximum for all fixed $1 \geq \epsilon > 0$ and $T > 0$, and Riemann-integration implies that

$$(1.11) \quad \text{vol}_{\mathbb{R}^d}(P_{\mathcal{S},\epsilon,T}) < \infty.$$

For Theorem 1.5 below, we will only consider the sets

$$P_{\mathcal{S},\epsilon,T} \quad \text{and} \quad P_{A,\epsilon,T}$$

for A with zero measure boundary contained in $\mathcal{S}(\delta)$ for some $\delta > 0$. For Theorem 1.6, we will consider some sets outside of \mathcal{S} .

1.5. Statement of results for lattice approximates. Let dk denote the probability Haar measure on $K := K_d := \text{SO}_d(\mathbb{R})$. Our main number-theoretic results are three averaged spiraling of lattice approximates results, one for approximation in the setting of Diophantine approximation with weights and two in the setting of multiplicative Diophantine approximation. We point out that our proof of Theorems 1.3 and 1.5 gives that the equality of the numerator and the equality of the denominator hold independently. One consequence is that other ratios may be obtained.

Theorem 1.3. *For every unimodular lattice $\Lambda \in X_d$, subset $A \subset \mathbb{S}^{m-1}$ with zero measure boundary, and $\epsilon > 0$, we have that*

$$\lim_{T \rightarrow \infty} \frac{\int_K \#\{k\Lambda \cap R_{A,\epsilon,T}\} dk}{\int_K \#\{k\Lambda \cap R_{\epsilon,T}\} dk} = \frac{\text{vol}_{\mathbb{R}^d}(R_{A,\epsilon,1})}{\text{vol}_{\mathbb{R}^d}(R_{\epsilon,1})}$$

REMARK 1.4. The special case of setting \mathbf{r} equal to $(1/m, \dots, 1/m)$ is, itself, already a generalization of [1, Theorem 1.4], except, since the function $\|\cdot\|_{(1/m, \dots, 1/m)}$ is (a power of) the sup norm, instead of the Euclidean norm of [1, Theorem 1.4]. Here, we obtain that the limit of the ratio is

$$\frac{\text{vol}_{\mathbb{R}^m}(R_{1,1} \cap \mathcal{CA})}{\text{vol}_{\mathbb{R}^m}(R_{1,1})},$$

where $\text{vol}_{\mathbb{R}^m}(R_{1,1}) = 2^m$. Note, as mentioned, the sets $R_{A,\epsilon,T}$ for the special case reduce to their counterparts in [1].

To obtain the exact generalization of [1, Theorem 1.4], replace the function $\|\cdot\|_{(1/m, \dots, 1/m)}$ by the Euclidean norm. Then the proof of the theorem will also give this generalization and the conclusion is that the limit of the ratios is $\text{vol}_{\mathbb{S}^{m-1}}(A)$. Note that, in all cases, the function $\|\cdot\|_{\mathbf{s}}$ can be for an arbitrary probability n -vector \mathbf{s} . We now state our results in the setting of multiplicative Diophantine approximation.

Theorem 1.5. *For every unimodular lattice $\Lambda \in X_d$, $\delta > 0$, subset $A \subset \mathcal{S}(\delta) =: \mathcal{S}$ with zero measure boundary, and $\epsilon > 0$, we have that*

$$\lim_{T \rightarrow \infty} \frac{\int_K \#\{k\Lambda \cap P_{A,\epsilon,T}\} dk}{\int_K \#\{k\Lambda \cap P_{\mathcal{S},\epsilon,T}\} dk} = \frac{\text{vol}_{\mathbb{R}^d}(P_{A,\epsilon,1})}{\text{vol}_{\mathbb{R}^d}(P_{\mathcal{S},\epsilon,1})}$$

Theorem 1.6. *For every unimodular lattice $\Lambda \in X_d$ and open subset $A \subset \mathbb{S}^{m-1}$ such that*

$$A \cap (\cup_{i=1}^m \mathbb{S}_i) \neq \emptyset,$$

we have that

$$\lim_{T \rightarrow \infty} \int_K \#\{k\Lambda \cap P_{A,\epsilon,T}\} dk = \infty.$$

Theorem 1.6 tells us that on average there are arbitrarily small neighborhoods of directions (which we know explicitly) for which every unimodular lattice has infinitely many elements in our star body. To

prove these theorems, we need our main ergodic result on equidistribution of Siegel translates, Theorem 2.2. We note that the spiraling results for multiplicative and weighted Diophantine approximation follow by applying the Theorems above to the unimodular lattice

$$\begin{pmatrix} \text{Id}_{m \times m} & \alpha \\ 0 & \text{Id}_{n \times n} \end{pmatrix} \mathbb{Z}^d$$

attached to a matrix $\alpha = (\alpha_{ij})$ as usual.

Acknowledgements. Part of this work was completed during the Group Actions and Number theory (GAN) programme at the Isaac Newton Institute for Mathematical Sciences. We thank the INI for providing a nice venue.

2. EQUIDISTRIBUTION ON THE SPACE OF LATTICES

Given a unimodular lattice Λ in \mathbb{R}^d and a bounded Riemann-integrable function f with compact support on \mathbb{R}^d , denote by \hat{f} its *Siegel transform*¹:

$$\hat{f}(\Lambda) := \sum_{\mathbf{v} \in \Lambda \setminus \{\mathbf{0}\}} f(\mathbf{v}).$$

Let $\mu = \mu_d$ be the probability measure on $X_d := \text{SL}_d(\mathbb{R})/\text{SL}_d(\mathbb{Z})$ induced by the Haar measure on $\text{SL}_d(\mathbb{R})$ and $d\mathbf{v}$ denote the usual volume measure on \mathbb{R}^d . We recall the classical Siegel Mean Value Theorem [22]:

Theorem 2.1. *Let f be as above.² Then $\hat{f} \in L^1(X_d, \mu)$ and*

$$\int_{\mathbb{R}^d} f \, d\mathbf{v} = \int_{X_d} \hat{f} \, d\mu.$$

Note that if f is the indicator function of a set $A \setminus \{\mathbf{0}\}$, then $\hat{f}(\Lambda)$ is simply the number of points in $\Lambda \cap (A \setminus \{\mathbf{0}\})$. Let

$$g_t := g_t^{(\mathbf{r}, \mathbf{s})} := \text{diag}(e^{r_1 t}, \dots, e^{r_m t}, e^{-s_1 t}, \dots, e^{-s_n t}) \in \text{SL}_d(\mathbb{R})$$

and e_1, \dots, e_d be the standard basis of \mathbb{R}^d . We use $\hat{\mathbf{1}}_A$ to denote the indicator function of the set A . Setting t so that $e^t = T$ gives

$$g_t R_{\epsilon, T} = R_{\epsilon, 1} =: R_{\epsilon} \quad \text{and} \quad g_t P_{\epsilon, T} = P_{\epsilon, 1} =: P_{\epsilon}$$

and

$$g_t R_{A, \epsilon, T} = R_{A, \epsilon, 1} =: R_{A, \epsilon} \quad \text{and} \quad g_t P_{A, \epsilon, T} = P_{A, \epsilon, 1} =: P_{A, \epsilon}.$$

Given a unimodular lattice $\Lambda \in \text{SL}_d(\mathbb{R})/\text{SL}_d(\mathbb{Z})$, a simple computation shows that

$$(2.1) \quad \#\{k\Lambda \cap R_{\epsilon, T}\} = \hat{\mathbf{1}}_{R_{\epsilon}}(g_t k\Lambda) \quad \text{and} \quad \#\{k\Lambda \cap P_{\epsilon, T}\} = \hat{\mathbf{1}}_{P_{\epsilon}}(g_t k\Lambda)$$

$$(2.2) \quad \#\{k\Lambda \cap R_{A, \epsilon, T}\} = \hat{\mathbf{1}}_{R_{A, \epsilon}}(g_t k\Lambda) \quad \text{and} \quad \#\{k\Lambda \cap P_{A, \epsilon, T}\} = \hat{\mathbf{1}}_{P_{A, \epsilon}}(g_t k\Lambda).$$

¹One could define the Siegel transform only over primitive lattice points, in which case results analogous to Theorems 2.2 and 2.3 also hold (using, essentially, the same proof).

²This condition can be generalized to $f \in L^1(\mathbb{R}^d)$.

2.1. Statement of results for Siegel transforms. To prove Theorems 1.3 and 1.5, we need to show the equidistribution of the Siegel transforms of the sets $R_{A,\epsilon}$, $P_{A,\epsilon}$, R_ϵ , and $P_{S,\epsilon}$ with respect to averages over $g_t^{(r,s)}$ -translates of K . The main ergodic tool in this setting is our fourth main theorem, a result on the mutiparameter spherical averages of Siegel transforms:

Theorem 2.2. *Let f be a bounded Riemann-integrable function of compact support on \mathbb{R}^d . Then for any $\Lambda \in X_d$,*

$$\lim_{t \rightarrow \infty} \int_{K_d} \widehat{f}(g_t^{(r,s)} k \Lambda) \, dk = \int_{X_d} \widehat{f} \, d\mu.$$

The above theorem is the generalization to the multiparameter case of our theorem for the single parameter case [1, Theorem 2.2]. Unlike in the single parameter case where the proof can be assembled from the work of Kleinbock-Margulis [11, Appendix], the multiparameter case cannot, as far as we are aware. Instead, we generalize our proof of [1, Theorem 2.2]. As in [1], the substantial part of the argument lies in the upper bound.

Theorem 2.3. *Let f be a bounded function of compact support in \mathbb{R}^d whose set of discontinuities has zero Lebesgue measure. Then for any $\Lambda \in X_d$,*

$$\lim_{t \rightarrow \infty} \int_{K_d} \widehat{f}(g_t^{(r,s)} k \Lambda) \, dk \leq \int_{X_d} \widehat{f} \, d\mu.$$

REMARK 2.4. The assumption that f has compact support can be replaced with that of $f \in L^1(\mathbb{R}^d)$ —the other assumptions are still, however, necessary for the proof.

Corollary 2.5. *Let f be a bounded Riemann-integrable function of compact support in \mathbb{R}^d . Then for any $\Lambda \in X_d$,*

$$\lim_{t \rightarrow \infty} \int_{K_d} \widehat{f}(g_t^{(r,s)} k \Lambda) \, dk \leq \int_{X_d} \widehat{f} \, d\mu.$$

Proof. Immediate from the theorem and the Lebesgue criterion. \square

As mentioned in [1], the lower bound follows either from the methods in [12] or by applying the following equidistribution theorem (Theorem 2.6) of Duke, Rudnick and Sarnak (cf. [6]) (see also Eskin and McMullen [9] and Shah [21]) and then approximating the Siegel transform \widehat{f} from below by $h \in C_c(X_d)$.

Theorem 2.6. *Let G be a non-compact semisimple Lie group and let K be a maximal compact subgroup of G . Let Γ be a lattice in G , let λ be the probability Haar measure on G/Γ , and let ν be any probability measure on K which is absolutely continuous with respect to a Haar measure on K . Let $\{a_n\}$ be a sequence of elements of G without accumulation points. Then for any $x \in G/\Gamma$ and any $h \in C_c(G/\Gamma)$,*

$$\lim_{n \rightarrow \infty} \int_K h(a_n k x) \, d\nu(k) = \int_{G/\Gamma} h \, d\lambda.$$

REMARK 2.7. One can replace dk by $d\nu(k)$ in Theorems 2.2 and 2.3 without any changes to the proofs.

2.2. Proof of Theorems 1.3 and 1.5. We prove Theorems 1.3 and 1.5 using Theorem 2.2, while deferring the proof of the latter to Section 3. Thus, applying Theorem 2.2 to the indicator function of $R_{A,\epsilon}$, we obtain

$$\lim_{t \rightarrow \infty} \int_K \widehat{\mathbf{1}}_{R_{A,\epsilon}}(g_t^{(r,s)} k \Lambda) dk = \int_{X_d} \widehat{\mathbf{1}}_{R_{A,\epsilon}} d\mu = \text{vol}_{\mathbb{R}^d}(R_{A,\epsilon}),$$

where we have applied Siegel's mean value theorem in the last equality.³ Doing likewise for R_ϵ , $P_{A,\epsilon}$, and $P_{S,\epsilon}$, we obtain

³A proof that $\widehat{\mathbf{1}}_{R_{A,\epsilon}}, \widehat{\mathbf{1}}_{R_\epsilon}, \widehat{\mathbf{1}}_{P_{A,\epsilon}}, \widehat{\mathbf{1}}_{P_{S,\epsilon}}$ are Riemann-integrable is analogous to that in [1, Footnote 4].

$$\begin{aligned} \lim_{T \rightarrow \infty} \frac{\int_K \# \{k\Lambda \cap R_{A,\epsilon,T}\} dk}{\int_K \# \{k\Lambda \cap R_{\epsilon,T}\} dk} &= \frac{\text{vol}_{\mathbb{R}^d}(R_{A,\epsilon})}{\text{vol}_{\mathbb{R}^d}(R_{\epsilon})} \\ \lim_{T \rightarrow \infty} \frac{\int_K \# \{k\Lambda \cap P_{A,\epsilon,T}\} dk}{\int_K \# \{k\Lambda \cap P_{S,\epsilon,T}\} dk} &= \frac{\text{vol}_{\mathbb{R}^d}(P_{A,\epsilon})}{\text{vol}_{\mathbb{R}^d}(P_{S,\epsilon})}, \end{aligned}$$

which proves our desired results. Note that (1.11) with $T = 1$ gives that $\text{vol}_{\mathbb{R}^d}(P_{S,\epsilon}) < \infty$.

2.3. Proof of Theorem 1.6. As in the Section 2.2, we use Theorem 2.2 before its proof. Let $\{\delta_i\}$ be a sequence of positive real numbers decreasing to 0. Then

$$A \supset \cup_i A \cap \mathcal{S}(\delta_i).$$

Let $\mathcal{C}_i := \mathcal{C}(A \cap \mathcal{S}(\delta_i))$. Applying Theorem 2.2, we have

$$\lim_{T \rightarrow \infty} \int_K \# \{k\Lambda \cap P_{A \cap \mathcal{S}(\delta_i), \epsilon, T}\} dk = \text{vol}_{\mathbb{R}^d}(P_{A \cap \mathcal{S}(\delta_i), \epsilon}) = O\left(\text{vol}_{\mathbb{R}^m}(\mathcal{C}_i)\right),$$

which $\rightarrow \infty$ as $i \rightarrow \infty$ by Lemma 4.1.

3. PROOF OF THEOREM 2.3

We adapt our proof in [1, Section 3] from the single parameter (i.e. the diagonal action is \mathbb{R} -rank 1) case to the multiparameter (i.e. the diagonal action is any allowed \mathbb{R} -rank) case. Recall our diagonal action is

$$g_t^{(\mathbf{r}, \mathbf{s})} =: g_t.$$

As mentioned, to prove Theorem 2.2, we need only show the upper bound (Theorem 2.3):

$$(3.1) \quad \lim_{t \rightarrow \infty} \int_{K_d} \widehat{f}(g_t k \Lambda) dk \leq \int_{X_d} \widehat{f} d\mu.$$

Fix a unimodular lattice $\Lambda \in X_d$. The strategy of the proof is to approximate using step functions on balls. We will divide the proof into four types of multiparameter actions:

- (1) $\mathbf{r} := (1/m, \dots, 1/m)$ and $n = 1$.
- (2) \mathbf{r} is an arbitrary probability m -vector and $n = 1$.
- (3) \mathbf{r} is an arbitrary probability m -vector and \mathbf{s} is an probability n -vector such there exist a unique entry j for which $s_j = \|\mathbf{s}\|$ where $\|\cdot\|$ is the sup norm.
- (4) \mathbf{r} is an arbitrary probability m -vector and \mathbf{s} is an arbitrary probability n -vector.

The first type is just our single parameter case [1, Theorem 2.2].

3.1. Proof for the second type of multiparameter. In this section, \mathbf{r} is an arbitrary probability m -vector and $n = 1$.

Using [1, Section 3.4] without change, we will approximate using step functions on balls, where we use the norm on $\mathbb{R}^d = \mathbb{R}^m \times \mathbb{R}$ given by the maximum of the Euclidean norm in $\mathbb{R}^m = \text{span}(e_1, \dots, e_m)$ and the absolute value in $\mathbb{R} = \text{span}(e_d)$. Hence, balls will be open regions of \mathbb{R}^d , which we also refer to as *rods* or *solid cylinders*. As in [1], we need four cases: balls centered at $\mathbf{0} \in \mathbb{R}^d$, balls centered in $\text{span}(e_d) \setminus \{\mathbf{0}\}$, balls centered in $\text{span}(e_1, \dots, e_m) \setminus \{\mathbf{0}\}$, and all other balls. Since we will approximate using step functions, it suffices (as we had shown in [1, Section 3.4]) to assume that the balls in the second case do not meet $\mathbf{0}$ and in the last case do not meet $\text{span}(e_d) \cup \text{span}(e_1, \dots, e_m)$.⁴ Let $E := B(\mathbf{w}, r)$ be any such ball and χ_E be its characteristic function. By the monotone convergence theorem, we have

$$\int_{K_d} \widehat{\chi}_E(g_t k \Lambda) dk = \sum_{\mathbf{v} \in \Lambda \setminus \{\mathbf{0}\}} \int_{K_d} \chi_{k^{-1} g_t^{-1} E}(\mathbf{v}) dk.$$

It is more convenient to prove the second and fourth cases together and before the others. Let E be a rod in either of these two cases. Let r be small. Let $R := e^t$. Fix R , or equivalently, t be a large

⁴We note that the second and the fourth cases already suffice to show Theorems 1.3, 1.5, and 1.6.

value. Now $g_t^{-1}E$ is also a rod, but narrow in the directions given by \mathbb{R}^m and long in the direction given by e_d . Recall from [1, Section 3], we have

$$(3.2) \quad \int_{K_d} \chi_{k^{-1}g_t^{-1}E}(\mathbf{v}) \, dk =: A_R^E(\|\mathbf{v}\|)$$

and

$$(3.3) \quad A_R^E(\tau) = \frac{\text{vol}_{\tau\mathbb{S}^m}(\tau\mathbb{S}^m \cap g_t^{-1}E)}{\text{vol}_{\tau\mathbb{S}^m}(\tau\mathbb{S}^m)}.$$

Also, recall, from [1, Section 3], the definition of a *cap* $\mathfrak{C}(\tau)$, namely it is the intersection of the rod $g_t^{-1}E$ with the sphere $\tau\mathbb{S}^m$. Now, unlike in [1], the caps are no longer spherical, but, for fixed R , are ellipsoidal of fixed eccentricity. All our geometric considerations are for a fixed R (which is only allowed to $\rightarrow \infty$ at the end). In particular, $A_R^E(\tau)$ is a strictly decreasing smooth function with respect to τ . Let $B_{\text{Euc}}(\mathbf{0}, \tau)$ denote a ball of radius τ in \mathbb{R}^d with respect to the Euclidean norm. Now it follows from the formula for A_R^E that

$$\sum_{\mathbf{v} \in \Lambda \setminus \{\mathbf{0}\}} A_R^E(\|\mathbf{v}\|) \leq \int_{\tau_-}^{\tau_+} \#(B_{\text{Euc}}(\mathbf{0}, \tau) \cap \Lambda \setminus \{\mathbf{0}\}) \, (-dA_R^E(\tau))$$

where the integral is the Riemann-Stieltjes integral and the integrability of the function $\#(B_{\text{Euc}}(\mathbf{0}, \tau) \cap \Lambda \setminus \{\mathbf{0}\})$ follows from its monotonicity and the continuity and monotonicity of $A_R^E(\tau)$. The rest of the proof is identical to that in [1, Section 3] and shows

$$\lim_{t \rightarrow \infty} \int_{K_d} \widehat{\chi}_E(g_t k \Lambda) \, dk \leq \text{vol}_{\mathbb{R}^d}(E).$$

Finally, we prove the first and third case together. Let E be a rod in either of these two cases. The difference between these two cases and the second and fourth cases is that the rod extends in both the positive e_d and negative e_d directions. As the lattice Λ is fixed, there is a ball $B_{\text{Euc}}(\mathbf{0}, \tau_0)$ in \mathbb{R}^d that does not meet $\Lambda \setminus \{\mathbf{0}\}$ for some $\tau_0 > 0$ depending only on Λ . Therefore, we can consider the two ends separately. The proof is the same as in [1, Section 3.3], except that \mathfrak{B} is not a sphere, but an ellipsoid of fixed eccentricity depending on R (which, recall is fixed until the end of the proof), but this does not affect the proof. Consequently, for the second type of multiparameter, we can conclude

$$\lim_{t \rightarrow \infty} \int_{K_d} \widehat{\chi}_E(g_t k \Lambda) \, dk \leq \text{vol}_{\mathbb{R}^d}(E).$$

3.2. Proof for the third type of multiparameter. In this section, \mathbf{r} is an arbitrary probability m -vector and \mathbf{s} is a probability n -vector such there exist a unique entry j for which $s_j = \|\mathbf{s}\|$ where $\|\cdot\|$ is the sup norm. On the other hand, for the rods that we define for this multiparameter type, we will use the norm on $\mathbb{R}^d = \mathbb{R}^m \times \mathbb{R}^n$ given by the maximum of the Euclidean norm in $\text{span}(e_1, \dots, e_{m+j-1}, e_{m+j+1}, \dots, e_d)$ and the absolute value in $\text{span}(e_{m+j})$. As before, we have four cases: balls centered at $\mathbf{0} \in \mathbb{R}^d$, balls centered in $\text{span}(e_{m+j}) \setminus \{\mathbf{0}\}$, balls centered in $\text{span}(e_1, \dots, e_{m+j-1}, e_{m+j+1}, \dots, e_d) \setminus \{\mathbf{0}\}$, and all other balls (again, Footnote 4 applies). Again, we may assume that the balls in the second case do not meet $\mathbf{0}$ and in the last case do not meet $\text{span}(e_{m+j}) \cup \text{span}(e_1, \dots, e_{m+j-1}, e_{m+j+1}, \dots, e_d)$.

Now g_t^{-1} has a unique largest expanding direction, namely along e_{m+j} . Replace the role of e_d from Section 3.1 with e_{m+j} . Let $R = e^{s_j t}$. Fix a large R , then the analysis of the geometry of $g_t^{-1}E$ is analogous to that in Section 3.1 because, for a fixed large R , the rod is much longer in along the e_{m+j} direction than any other. The only difference is that there exists a minimum sphere radius $\tilde{\tau}(R)$ larger than which the analysis of the geometry is valid because some directions are expanding (but less than in e_{m+j}). However, for R large, $\tilde{\tau}(R)$ is small in comparison to the length of the rod $\tau_+(R)$ (which is on the order of R). In particular, $\lim_{R \rightarrow \infty} \tilde{\tau}(R)/\tau_+(R) = 0$. Consequently (as shown in [1,

Section 3.3] for example, the error is $O(R^{-1})$, which does not affect the proof. The conclusion, in all four cases, is

$$\lim_{t \rightarrow \infty} \int_{K_d} \hat{\chi}_E(g_t k \Lambda) \, dk \leq \text{vol}_{\mathbb{R}^d}(E).$$

3.3. Proof for the fourth type of multiparameter. In this section, \mathbf{r} is an arbitrary probability m -vector and \mathbf{s} is an arbitrary probability n -vector. We may assume without loss of generality that there exist indices $1 \leq j_1 < \dots < j_\ell \leq n$ such that $s_{j_1} = \dots = s_{j_\ell} = \|\mathbf{s}\| =: \lambda$ and $2 \leq \ell \leq n$. (Again $\|\cdot\|$ denotes the sup norm.) Let $\tilde{\lambda}$ denote the largest component of \mathbf{s} strictly less than λ , or, if no such component exists, set $\tilde{\lambda} = 1$. Let us denote this set of indices J and the remaining indices by J^c , and note that $J \sqcup J^c = \{1, \dots, n\}$. The main difference and problem with this case is that caps are no longer relatively small in relation to the largest dimension of the rod. To take care of this problem, we adapt the proof in Section 3.2 in two ways, the first for the analog of first and third cases and the second for the analog of the second and fourth cases.

We use two types of balls/rods. For the balls/rods that we define for this multiparameter type for the first and third case, we will use the norm on $\mathbb{R}^d = \mathbb{R}^m \times \mathbb{R}^n$ given by the maximum of the Euclidean norm in

$$\text{span}\left(\bigcup_{i=1}^m e_i \cup \bigcup_{j \in J^c} e_{m+j}\right)$$

and the sup norm in

$$\text{span}\left(\bigcup_{j \in J} e_{m+j}\right).$$

For the balls/rods that we define for this multiparameter for the second and fourth cases, we use the sup norm until almost the end of the proof (again, Footnote 4 applies). As before, we have the four cases: balls centered at $\mathbf{0} \in \mathbb{R}^d$, balls centered in $\text{span}(\bigcup_{j \in J} e_{m+j}) \setminus \{\mathbf{0}\}$, balls centered in $\text{span}(\bigcup_{i=1}^m e_i \cup \bigcup_{j \in J^c} e_{m+j}) \setminus \{\mathbf{0}\}$, and all other balls. Again, we may assume that the balls in the second case do not meet $\mathbf{0}$ and in the last case do not meet $\text{span}(\bigcup_{i=1}^m e_i \cup \bigcup_{j \in J^c} e_{m+j}) \cup \text{span}(\bigcup_{j \in J} e_{m+j})$. Let $E := B(\mathbf{w}, r)$. We prove each case in turn—for convenience of exposition, we prove the cases in the order first, third, second, and fourth.

3.3.1. The first case: balls centered at $\mathbf{0}$. Let $R = e^{\lambda t}$. Fix a large R . Consider the rod $g_t^{-1}E$. The directions J are all expanded to a radius of Rr . All other directions are expanding less or contracting. As in Section 3.2, there exists a minimal radius $\tilde{\tau}(R)$ larger than which the analysis of the geometry is valid and we can choose $\tilde{\tau}(R) = 3e^{\tilde{\lambda}t}$; hence, we have $\lim_{R \rightarrow \infty} \tilde{\tau}(R)/Rr = 0$, which implies we can ignore radius smaller than $\tilde{\tau}(R)$. As mentioned, caps $\mathfrak{C}(\tau)$ are no longer small, but this does not affect the analysis of the geometry from Section 3.2 up to the inequality

$$\begin{aligned} \sum_{\mathbf{v} \in \Lambda \setminus \{\mathbf{0}\}} A_R^E(\|\mathbf{v}\|) &\leq O(R^{-\ell}) + d \int_{\tilde{\tau}}^{Rr} (1 + \varepsilon) \text{vol}(B_{\text{Euc}}(\mathbf{0}, 1)) \frac{C(\tilde{\tau})}{\text{vol}_{\mathbb{S}^d}(\mathbb{S}^d)} \, d\tau \\ &= O(R^{-\ell}) + d(1 + \varepsilon) \frac{\text{vol}(B_{\text{Euc}}(\mathbf{0}, 1))}{\text{vol}_{\mathbb{S}^d}(\mathbb{S}^d)} C(\tilde{\tau})(Rr - \tilde{\tau}) \end{aligned}$$

where $C(\tau) = \text{vol}_{\mathbb{S}^d} \mathfrak{C}(\tau)$ and the $O(R^{-\ell})$ comes from $\tau < \tilde{\tau}(R)$. Now $C(\tilde{\tau})(Rr - \tilde{\tau})$ is the volume of the rod with a (relatively) small hole missing. Letting $R \rightarrow \infty$ and $\varepsilon \rightarrow 0$ yields our desired result:

$$\lim_{t \rightarrow \infty} \int_{K_d} \hat{\chi}_E(g_t k \Lambda) \, dk \leq \text{vol}_{\mathbb{R}^d}(E).$$

3.3.2. *The third case: balls centered at $\text{span}(\bigcup_{i=1}^m e_i \cup \bigcup_{j \in J^c} e_{m+j}) \setminus \{\mathbf{0}\}$.* The proof is similar to the first case. In any expanding directions $\bigcup_{j \in J} e_{m+j}$, the components of \mathbf{w} are zero and hence $g_t^{-1}\mathbf{w}$ has at most expansion at a rate of $e^{\tilde{\lambda}t}$. Let $R = e^{\lambda t}$. Fix a large R . Let $\tilde{\varepsilon} := \tilde{\varepsilon}(R) := \|g_t^{-1}\mathbf{w}\|/R$. Hence, $\lim_{R \rightarrow \infty} \tilde{\varepsilon}(R) = 0$. Consequently, using the analogous proof as in the first case for a slightly larger rod (replacing r with $(1 + 2\tilde{\varepsilon})r$), we have

$$\sum_{\mathbf{v} \in \Lambda \setminus \{\mathbf{0}\}} A_R^E(\|\mathbf{v}\|) \leq O(R^{-\ell}) + d \int_{\tilde{\tau}}^{\tau_+} (1 + \varepsilon) \text{vol}(B_{\text{Euc}}(\mathbf{0}, 1)) \frac{C(\tilde{\tau})}{\text{vol}_{\mathbb{S}^d}(\mathbb{S}^d)} d\tau$$

where $\tau_+ := Rr(1 + 2\tilde{\varepsilon})$ and $C(\tilde{\tau})(\tau^+ - \tilde{\tau}) \rightarrow \text{vol}_{\mathbb{R}^d}(E)$ as $R \rightarrow \infty$. This yields our desired result:

$$\lim_{t \rightarrow \infty} \int_{K_d} \hat{\chi}_E(g_t k \Lambda) dk \leq \text{vol}_{\mathbb{R}^d}(E).$$

3.3.3. *The second case: balls centered in $\text{span}(\bigcup_{j \in J} e_{m+j}) \setminus \{\mathbf{0}\}$.* Of the indices in J pick one, say j_1 . Let us first consider the special case that $\mathbf{w} = w e_{m+j_1}$ for some $w \neq 0$. This index will play the role of d from Section 3.1. Let $I = \{1, \dots, m\} \cup \{m+j : j \in J^c\}$ and $R = e^{\lambda t}$. For the second (and fourth) cases, we will assume an additional condition (which we later show does not affect the generality of our result): for a fixed $\alpha \geq 1$, we only consider balls E for which

$$(3.4) \quad \frac{\text{dist}(\mathbf{w}, \text{span}(\bigcup_{i \in I} e_i)) - r}{r} \geq \alpha$$

holds. Recall that our ball E is given by the sup norm—it is a d -cube. Its translate $E - \mathbf{w}$ has exactly one vertex \mathbf{p} with all positive coordinates. Let us change E into a “half-closed” ball F by the union of all the $d - 1$ hyperfaces of the cube with $\mathbf{p} + \mathbf{w}$ as vertex. Any half-closed ball will be constructed like this. We will refer to F and its g_t -translates as *half-closed rods* or simply *rods* if the context is clear. Fix a large R . Consider the rod $g_t^{-1}F$ and one of the $d - 1$ -dimensional faces that is normal to e_{m+j_1} —call this face \mathcal{F} and note that it is a $d - 1$ -dimensional box.

Choose a large natural number N . Partition the smallest side length of \mathcal{F} into N segments of length L . For each of the other side lengths in \mathcal{F} , partition into segments whose length is nearest to L . This partitions \mathcal{F} into $\mathcal{N}(N)$ boxes with the same side lengths and, furthermore, each of which can be contained in a $d - 1$ -dimensional cube of side length $2L < 2/N$. Let us index these little boxes by k . The cartesian product of each of these little cubes with the e_{m+j_1} -th coordinate of $g_t^{-1}F$ are rods, which we make into half-closed rods in the way specified above. This is a partition of $g_t^{-1}F$ such that there is only one direction, namely e_{m+j_1} , that is long. To each element of this partition, cases two and four of Section 3.1 applies (the fact that each element is a half-closed rod as opposed to an open rod does not affect the proof in Section 3.1). Since this is a partition, elements are pairwise disjoint and we may sum over each element of the partition to obtain

$$\begin{aligned} \sum_{\mathbf{v} \in \Lambda \setminus \{\mathbf{0}\}} A_R^F(\|\mathbf{v}\|) &\leq d \sum_{k=1}^{\mathcal{N}(N)} \int_{\tau_k^-}^{\tau_k^+} (1 + \varepsilon) \text{vol}(B_{\text{Euc}}(\mathbf{0}, 1)) \frac{C_k(\tau_k^-)}{\text{vol}_{\mathbb{S}^d}(\mathbb{S}^d)} d\tau \\ &= d(1 + \varepsilon) \frac{\text{vol}(B_{\text{Euc}}(\mathbf{0}, 1))}{\text{vol}_{\mathbb{S}^d}(\mathbb{S}^d)} \sum_{k=1}^{\mathcal{N}(N)} C_k(\tau_k^-)(\tau_k^+ - \tau_k^-) \end{aligned}$$

where τ_k^- and τ_k^+ are, up to $O(R^{-1})$, the minimum and maximum radii such that $\tau \mathbb{S}^d$ meets the k -th partition element and $C_k(\tau)$ is the volume of the cap of the k -th element, i.e. $C_k(\tau) = \text{vol}_{\tau \mathbb{S}^d}(\mathfrak{C}_k(\tau))$ where $\mathfrak{C}_k(\tau)$ is the intersection of k -th partition element with $\tau \mathbb{S}^d$. Within $O(R^{-1})$, $\tau_k^+ - \tau_k^-$ is the length of the rod $g_t^{-1}F$ along the e_{m+j_1} direction. Now $C_k(\tau_-)(\tau_+ - \tau_-)$ is the volume of an element that has length along e_{m+j_1} within $O(R^{-1})$ of the length along e_{m+j_1} of $g_t^{-1}F$, but with cross-section volume $C_k(\tau_-)$. Since in the second case (3.4) holds, a direct calculation (using trigonometry) gives

that

$$(3.5) \quad C_k(\tau_-) \leq \tilde{\gamma} \frac{\text{vol}_{\mathbb{R}^{d-1}}(\mathfrak{B}_k)}{\sin^{d-1}(\pi/2 - \arcsin(1/\alpha))}$$

where $\tilde{\gamma} > 1$ is a number depending only N and α such that $\tilde{\gamma} \searrow 1$ as $N, \alpha \rightarrow \infty$ and \mathfrak{B}_k is the interesection of the $d - 1$ -dimensional hyperplane normal to e_{m+j_1} with the k -th partition element. Consequently, for large N, α ,

$$\sum_{k=1}^{\mathcal{N}(N)} C_k(\tau_k^-)(\tau_k^+ - \tau_k^-)$$

is arbitrarily close to $\text{vol}_{\mathbb{R}^d}(E)$. As $\frac{\text{vol}(B_{\text{Euc}}(\mathbf{0}, 1))}{\text{vol}_{\mathbb{S}^d}(\mathbb{S}^d)} = \frac{1}{d+1}$, letting $R \rightarrow \infty$ and $\varepsilon \rightarrow 0$, we have our desired result for the special case:

$$\lim_{t \rightarrow \infty} \int_{K_d} \hat{\chi}_E(g_t k \Lambda) \, dk \leq \text{vol}_{\mathbb{R}^d}(E),$$

up to the restrictions that the balls are now half-closed and that (3.4) must hold. Likewise, we have the same conclusion for $\mathbf{w} = w e_{m+j}$ for any $j \in J$.

We now consider the second case in general. We may assume that $\mathbf{w} \in \text{span}(\bigcup_{j \in J} e_{m+j}) =: \mathbb{P}_J$ but not in any of the coordinate axes. Let $\mathbf{q} := \mathbf{q}(R)$ denote the point of $\overline{g_t^{-1}F}$ with smallest Euclidean norm. By convexity, it is easy to see that the point \mathbf{q} is unique (for fixed R) and that $\mathbf{q} \in \text{span}(\bigcup_{j \in J} e_{m+j})$. We remark that \mathbf{q} is an eigenvector of g_t^{-1} and thus the direction of \mathbf{q} is fixed for all R . Let $\|\cdot\|_J$ be the sup norm in \mathbb{P}_J . Rotate a coordinate axis to the direction of \mathbf{q} —doing this to a half-closed $\|\cdot\|_J$ -ball of radius β yields a rotated half-closed ℓ -cube $\mathcal{C}(\beta)$ with side length 2β . Cover $F \cap \mathbb{P}_J$ by a partition of affine translates of $\bigcup_t \mathcal{C}_t(\beta)$ where $\beta > 0$ is a constant so small that $\text{vol}_{\mathbb{R}^d}(F \cap \mathbb{P}_J)$ is less than but as close to $\text{vol}_{\mathbb{R}^d}(\bigcup_t \mathcal{C}_t(\beta))$ as desired. For each $\mathcal{C}_t(\beta)$ take the cartesian product with the other directions of F to obtain $\tilde{F}_t(\beta)$. Then $\text{vol}_{\mathbb{R}^d}(F)$ is less than but as close to $\text{vol}_{\mathbb{R}^d}(\bigcup_t \tilde{F}_t(\beta))$ as desired. Choose t so large that $e^{\lambda t} \beta$ is larger than the R chosen in the special case (where the center is on the axis) above—this gives us a much larger R for this, the second case in general. Now the special case holds for each $\tilde{F}_t(\beta)$ and applying it to each and summing over the partition and noting that the volume of the partition is arbitrarily close to $\text{vol}_{\mathbb{R}^d}(E)$ yields the second case in general:

$$\lim_{t \rightarrow \infty} \int_{K_d} \hat{\chi}_E(g_t k \Lambda) \, dk \leq \text{vol}_{\mathbb{R}^d}(E),$$

up to the restrictions that the balls are now half-closed and that (3.4) must hold.

3.3.4. The fourth case: all other balls. This is an adaption of the second case. The difference is that $\mathbf{q} \notin \mathbb{P}_J$. Let $\|\cdot\|_I$ be the sup norm in the $\text{span}(\bigcup_{i \in I} e_i) =: \mathbb{P}_I$ and let \mathbf{q}_I and \mathbf{q}_J be the orthogonal projections of \mathbf{q} onto \mathbb{P}_I and \mathbb{P}_J , respectively. Then

$$\frac{\|\mathbf{q}_I(R)\|_I}{\|\mathbf{q}_J(R)\|_J} = 0$$

as $R \rightarrow \infty$. Consequently, for R large enough, (3.5) holds and thus the proof of the second case also applies to this case, allowing us to conclude:

$$\lim_{t \rightarrow \infty} \int_{K_d} \hat{\chi}_E(g_t k \Lambda) \, dk \leq \text{vol}_{\mathbb{R}^d}(E),$$

up to the restrictions that the balls are now half-closed and that (3.4) must hold.

3.3.5. Finishing the second and fourth cases. We wish to prove the second and fourth cases for the balls defined for the first and third cases (i.e. in terms of the product of Euclidean norms). To remove the restriction of half-closed rods, consider the measure zero boundaries of the half-closed rods at each stage. Using [1, Lemma 3.5] to approximate this measure zero set and the method of handling the null term from [1, Section 3.4], we can remove this restriction. To remove the restriction given by (3.4), we note that [1, Lemmas 3.1 and 3.5] apply to the balls of the second and fourth case with the restriction (3.4) because α is fixed and the ball given by the product of the Euclidean norms not do meet \mathbb{P}_I . This is all that is needed to apply [1, Section 3.4]. Doing so allows us to conclude

$$\lim_{t \rightarrow \infty} \int_{K_d} \hat{\chi}_E(g_t k \Lambda) \, dk \leq \text{vol}_{\mathbb{R}^d}(E),$$

where E is a ball in the same norm as for the first and third case.

3.4. Finishing the proof of Theorem 2.3. For each mutiparameter type, we apply [1, Section 3.4] without change.

4. APPENDIX

We prove Lemma 4.1. Recall that $\|\mathbf{v}\|_{\text{pr}} := \prod_{i=1}^{\ell} |v_i|$; let $\ell = m$ in this section. Let

$$S := \{\mathbf{v} \in \mathbb{R}^m : \|\mathbf{v}\|_{\text{pr}} \leq 1\}.$$

Lemma 4.1. *Let \mathbf{w} be on a great sphere for \mathbb{S}^{m-1} . Let $A := B(\mathbf{w}, r) \cap \mathbb{S}^{m-1}$ for some $r > 0$. Then*

$$\text{vol}_{\mathbb{R}^m}(\mathcal{C}A \cap S) = \infty.$$

Proof. For $m = 1$, a great sphere is simply the intersection of an axis with the circle. Elementary calculus gives the result.

We may assume that $m \geq 2$. Without loss of generality, we may assume that r is small. There are two cases. First assume that \overline{A} does not meet any coordinate axes. Then there exists exactly one coordinate in which the points in \overline{A} may have small absolute value. By reordering indices if necessary, we may assume that the m -th coordinate is the one that has small absolute values. In the other directions, the absolute values are bounded away from 0. In other words, given a constant $c > 0$, we have

$$\prod_{i=1}^{m-1} |v_i| \geq c$$

for all $\mathbf{v} \in \overline{A}$. Note that $\text{vol}_{\mathbb{S}^{m-1}}(A) = O(r^{m-1})$ and that

$$\prod_{i=1}^{m-1} |v_i| = O(\text{vol}_{\mathbb{S}^{m-1}}(A))$$

for all $\mathbf{v} \in \overline{A}$ (because r is small). Consequently, for large τ , we have that

$$\prod_{i=1}^{m-1} |\tau v_i| = O(\text{vol}_{\tau \mathbb{S}^{m-1}}(\tau A)).$$

Perhaps by cutting off the part of the cone nearest to the origin, we have that $\mathcal{C}A \cap S$ is the graph of the function over A determined by $\prod_{i=1}^m |x_i| = 1$, giving us that $\text{vol}_{\tau \mathbb{S}^{m-1}}(\tau A \cap S) |\tau v_m| = O(1)$. This implies that $|v_m| \tau^m \leq O(1)$. Riemann integration now gives

$$\text{vol}_{\mathbb{R}^m}(\mathcal{C}A \cap S) = \text{const} \int_1^\infty \text{vol}_{\tau \mathbb{S}^{m-1}}(\tau A \cap S) \, d\tau = \text{const} \int_1^\infty \frac{1}{|\tau v_m|} \, d\tau \geq \text{const} \int_1^\infty \tau^{m-1} \, d\tau = \infty.$$

We note that const depends on how close A is to a coordinate axis.

The other case is when \overline{A} meets coordinate axes. Since r is small, it may only meet one. Pick an open ball $\tilde{B} \subset A$ that avoids the axis and apply the previous proof. \square

5. CONCLUDING REMARKS AND GENERALISATIONS

As we have noted in [1], Theorem 2.2 is reminiscent of some results in the literature, for instance the work of Eskin-Margulis-Mozes, where the authors average over a different compact group. Closest to our work is the result of Kleinbock-Margulis who provide a proof in [11] Corollary A.8 of a very general averaging result for one parameter flows, or more precisely, flows with the property “EM” on quotients of semisimple groups by irreducible lattices. We refer the reader to [11] for precise statements and the definition of the property “EM”. The main tool in loc. cit. is exponential decay of correlation coefficients of Hölder vectors. In [1], we provide a new, elementary proof of the above theorem for one parameter flows on $\mathrm{SL}_n(\mathbb{R})/\mathrm{SL}_n(Z)$. On the other hand, as we have mentioned, the higher rank averaging result in the present paper is new.

Several authors have investigated variations of Dirichlet’s theorem in the context of number fields. Let k be a number field which we assume to be totally real for convenience, let S be the set of infinite places of k and let O_S be the ring of S -integers of k . For $x \in k^n$, we define the S height to be

$$h_S(x) := \prod_{v \in S} \max_{1 \leq i \leq n} \{|x_i|_v\}$$

where $|\cdot|_v$ denotes the v -adic valuation on the completion k_v of k . In [4], Burger proves⁵ that for every $x_v \in k_v$ and $Q > c(k)^{(1+n)/n}$, there exist $q \in O_S^n$ and $p \in O_S$ such that $h_S(q) \leq Q$ and

$$\prod_{v \in S} |x_v q - p|_v \leq c(k)^{1+n} Q^{-n}$$

where $c(k) = \Delta_k^{1/2d}$, d is the degree of the number field and Δ_k is the discriminant of k . Using the methods of [1] and the above mentioned Theorem of Kleinbock-Margulis applied to

$$G = \prod_{v \in S} \mathrm{SL}_{n+1}(k_v), \Gamma = \mathrm{SL}_{n+1}(O_S)$$

and $g_t = (\mathrm{diag}(e^{nt}, e^{-t}, \dots, e^{-t}))_{v \in S}$ one can obtain equidistribution of approximates for the analogue of Dirichlet’s theorem for number fields as above. Other generalisations of Dirichlet’s theorem have been established by W. M. Schmid[19], Qûeme [17], Hattori [10], see also [7] where another analogue of Dirichlet’s theorem and an analogue of badly approximable vectors in number fields is investigated. It would not be difficult to obtain equidistribution of approximates in each of these cases with appropriate choices of G , Γ and g_t . However, multiplicative, weighted and versions for more places (p -adic) do not follow from the above methods. It would be an interesting problem to generalise our higher rank result in this paper to include other semisimple Lie groups and also to include finite, i.e. p -adic places.

REFERENCES

- [1] Jayadev S. Athreya, Anish Ghosh, and Jimmy Tseng, *Spherical averages of Siegel transforms and spiraling of lattice approximations*, to appear in J. London Math Soc.
- [2] D. Badziahin, A. Pollington, S. Velani, *On a problem in simultaneous Diophantine approximation: Schmidt’s conjecture*, Ann. of Math. 174 (2011), 1837–1883.
- [3] Yann Bugeaud, *Multiplicative Diophantine approximation*, Dynamical systems and Diophantine approximation, 105–125, Semin. Congr., 19, Soc. Math. France, Paris, 2009.
- [4] E. Burger, *Homogeneous Diophantine approximation in S -integers*, Pacific J. Math. 152 (1992), no. 2, 211–253.
- [5] L. G. P. Dirichlet, *Verallgemeinerung eines Satzes aus der Lehre von den Kettenbrüchen nebst einige Anwendungen auf die Theorie der Zahlen*, S.-B. Preuss. Akad. Wiss. (1842), 93–95.
- [6] W. Duke, Z. Rudnick, and P. Sarnak. Density of integer points on affine homogeneous varieties. Duke Math. J., 71(1):143–179, 1993.
- [7] M. Einsiedler, A. Ghosh and B. Lytle, *Badly approximable vectors, C^1 curves, and number fields*, to appear in Ergodic Theory and Dynamical Systems.
- [8] A. Eskin, G. Margulis and S. Mozes, *Upper bounds and asymptotics in a quantitative version of the Oppenheim conjecture*, Ann. of Math. 147 (1998), no. 1, 93–141.
- [9] A. Eskin and C. McMullen, *Mixing, counting, and equidistribution in Lie groups*, Duke Math. J., 71(1):181–209, 1993.

⁵He actually proves his result for matrices with entries in k_v

- [10] T. Hattori, *Some Diophantine approximation inequalities and products of hyperbolic spaces*, J. Math. Soc. Japan 59 (2007), no. 1, 239–264.
- [11] D. Kleinbock and G. Margulis, *Bounded orbits of nonquasiunipotent flows on homogeneous spaces*, Amer. Math. Soc. Transl.(1996), v. 171, 141–172.
- [12] D. Kleinbock and G. Margulis, *Logarithm laws for flows on homogeneous spaces*, Invent. Math., 138 (1999), 451–494.
- [13] D. Kleinbock and B. Weiss, *Modified Schmidt games and Diophantine approximation with weights*, Advances in Math. 223 (2010), 1276–1298.
- [14] D. Kleinbock and B. Weiss, *Modified Schmidt games and a conjecture of Margulis*, J. Mod. Dyn. 7 (2014) 429–460.
- [15] J. Marklof and A. Strömbergsson, *The Boltzmann-Grad limit of the periodic Lorentz gas*, Annals of Mathematics 174 (2011) 225–298.
- [16] J. Marklof and A. Strömbergsson, *The distribution of free path lengths in the periodic Lorentz gas and related lattice point problems*, Annals of Mathematics 172 (2010) 1949–2033.
- [17] R. Qûeme, *On Diophantine approximation by algebraic numbers of a given number field: a new generalization of Dirichlet approximation theorem*, Journées Arithmétiques, 1989 (Luminy, 1989). Astérisque No. 198-200 (1991), 273–283 (1992).
- [18] W. M. Schmidt, *Asymptotic formulae for point lattices of bounded determinant and subspaces of bounded height*, Duke Math J. 35 (1968), 327–339.
- [19] ———, *Simultaneous approximation to algebraic numbers by elements of a number field*, Monatsh. Math. 79 (1975), 55–66.
- [20] ———, *Diophantine approximation*, Lecture Notes in Mathematics, vol. 785, Springer-Verlag, Berlin, 1980.
- [21] Nimish Shah, *Limit distributions of expanding translates of certain orbits on homogeneous spaces*, Proc. Indian Acad. Sci., Math. Sci. 106 (1996) 105–125.
- [22] C. S. Siegel, *A mean value theorem in geometry of numbers*, Ann. Math. 46 (1945), 340–347.

J.S.A.: DEPARTMENT OF MATHEMATICS, UNIVERSITY OF ILLINOIS URBANA-CHAMPAIGN, 1409 W. GREEN STREET, URBANA, IL 61801, USA

E-mail address: jathreya@illinois.edu

A.G.: SCHOOL OF MATHEMATICS, TATA INSTITUTE OF FUNDAMENTAL RESEARCH, HOMI BHABHA ROAD, MUMBAI 400005 INDIA

E-mail address: ghosh@math.tifr.res.in

J.T.: SCHOOL OF MATHEMATICS, UNIVERSITY OF BRISTOL, UNIVERSITY WALK, BRISTOL, BS8 1TW UK

E-mail address: j.tseng@bristol.ac.uk